UNITED STATES PATENT AND TRADEMARK OFFICE

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
09/892,884	06/28/2001	Tsutomu Abe	109967	3031

25944	7590	02/24/2005

OLIFF & BERRIDGE, PLC
P.O. BOX 19928
ALEXANDRIA, VA 22320

EXAMINER
SIANGCHIN, KEVIN

ART UNIT	PAPER NUMBER
2623	

DATE MAILED: 02/24/2005

Please find below and/or attached an Office communication concerning this application or proceeding.

PTO-90C (Rev. 10/03)

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --

Period for Reply

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE <u>3</u> MONTH(S) FROM
THE MAILING DATE OF THIS COMMUNICATION.
- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed
  after SIX (6) MONTHS from the mailing date of this communication.
- If the period for reply specified above is less than thirty (30) days, a reply within the statutory minimum of thirty (30) days will be considered timely.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133).
  Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any
  earned patent term adjustment. See 37 CFR 1.704(b).

Status

1)☒ Responsive to communication(s) filed on <u>07 September 2004</u>.

2a)☒ This action is **FINAL.**   2b)☐ This action is non-final.

3)☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is
closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

Disposition of Claims

4)☒ Claim(s) <u>1-4,6-12 and 14-40</u> is/are pending in the application.

    4a) Of the above claim(s) _____ is/are withdrawn from consideration.

5)☐ Claim(s) _____ is/are allowed.

6)☒ Claim(s) <u>1-4,6-12 and 14-40</u> is/are rejected.

7)☐ Claim(s) _____ is/are objected to.

8)☐ Claim(s) _____ are subject to restriction and/or election requirement.

Application Papers

9)☒ The specification is objected to by the Examiner.

10)☒ The drawing(s) filed on <u>14 September 2001</u> is/are: a)☒ accepted or b)☐ objected to by the Examiner.

    Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).

    Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).

11)☐ The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

Priority under 35 U.S.C. § 119

12)☒ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).

    a)☒ All b)☐ Some * c)☐ None of:

        1.☐ Certified copies of the priority documents have been received.

        2.☐ Certified copies of the priority documents have been received in Application No. _____.

        3.☐ Copies of the certified copies of the priority documents have been received in this National Stage
application from the International Bureau (PCT Rule 17.2(a)).

    * See the attached detailed Office action for a list of the certified copies not received.

Attachment(s)

1)☒ Notice of References Cited (PTO-892)

2)☐ Notice of Draftsperson's Patent Drawing Review (PTO-948)

3)☒ Information Disclosure Statement(s) (PTO-1449 or PTO/SB/08)
Paper No(s)/Mail Date <u>09/14/2001</u>.

4)☐ Interview Summary (PTO-413)
Paper No(s)/Mail Date. _____ .

5)☐ Notice of Informal Patent Application (PTO-152)

6)☐ Other: _____ .

Detailed Action

Drawings

Response to Drawing Corrections

1.   New drawings were received on 14th of September 2001.

2.   In the previous Office Action, the Examiner failed to acknowledge these drawings. The new drawings are acceptable. Consequently, all drawing objections of the previous Office Action are withdrawn.

3.   Issues may remain with Figs. 9 and 10. Refer to the comments below.

Specification

Response to Amendments to the Specification

4.   The amendments to the Specification, filed on the 7th of September 2004 (hereinafter, *Amendments to the Specification*, hereinafter) have been acknowledged.

5.   As a result of these amendments, the objections posed in items (a), (b), and (c) in paragraph 3 of the previous Office Action are overcome. The objections posed in paragraph 4 of the previous Office Action are likewise overcome.

6.   The Applicant has amended the Title of the Invention. However, the amended title still fails to adequately describe the Applicant's claimed invention. According to 37 C.F.R. § 1.72, the "title of the invention may not exceed 500 characters in length and must be as short and *specific* as possible"[1]. The Applicant's amended title will be treated in further detail below.

---

[1]   Emphasis added.

7.   With regard to Fig. 10 and its description in the Specification, the Applicant rightly observes: "Fig. 10

represents a slit pattern of the board 606, which has been newly coded" (Amendments to the Specification, page 18,

last paragraph). This, however, does not sufficiently clarify how one would arrive at the depicted pattern, having

contemplated the Applicant's disclosed methodology in light of the corresponding patterns shown in Figs. 8 and 9.

The confusion (at least on the part of the Examiner) arises less from a deficiency in the Applicant's disclosure, than

from deficiencies or errors in the accompanying figures (viz. Figs. 9 and 10). If, on the other hand, the drawings are

correct, then the Specification is indeed deficient, as pointed out in the previous Office Action.

8.   The scene in Fig. 6 is illuminated by the projected stripe pattern. Fig. 8 depicts this illuminated scene, as

viewed from the left most camera, camera 601 (cf. Fig. 6). In Fig. 8, notice that the stripe patterns of the square

region corresponding to box 606 proceed in the following order: area 801 (corresponding to the left face of the box),

white (W), light crosshatch (LC), medium crosshatch (MC), white (W), medium crosshatch (MC), and white (W).

9.   Fig. 9 depicts an image of the scene obtained from the right most camera, camera 602 (cf. Fig. 6). There,

the stripe patterns of the square region corresponding to box 606 proceed in the following order: W, LC, MC, W,

LC, W, and area 901 (corresponding to the right face of the box). Notwithstanding the side areas 801 and 901, the

order of the stripe patterns should be the same in both Fig. 8 and Fig. 9 (assuming the entire front face of the box is

visible to both cameras[2]), since the figures depict the same scene. In Figs. 8 and 9, the order is clearly different.

10.   The question that then arises is which of the two orders is correct, or at least most consistent with the

Applicant's disclosure. It should be clear, upon inspection, that the order depicted in Fig. 8 is most consistent with

the background region (i.e. the region of Fig. 8 corresponding to wall 605). For example, in Fig. 8 the two MC

stripes correspond to two medium crosshatched regions of the background, whereas in Fig. 9, one LC stripe

corresponds to a lightly crosshatched stripe of the background, while the other LC stripe corresponds to a medium

crosshatched region of the background. Fig. 9 should, therefore, be changed in a manner that conforms to Fig. 8.

This could be accomplished, for example, by changing the order of the stripes in the aforesaid square region from

W, LC, MC, W, LC, W, and area 901 to W, LC, MC, W, MC, W, and area 901.

---

2   This is a reasonable assumption because the ultimate goal of using the structured light patterns is to obtain stereo correspondences between
        the images obtained by both cameras.

11.   Assuming the correct stripe pattern appears in both Figs. 8 and 9, the stripes on the face of box **606** would have the following order: W, LC, MC, W, MC, and W (as shown in Fig. 8). Fig. 10, which "represents a slit pattern of the board 606", should then have an analogous pattern. Again, the pattern should be consistent with the stripe pattern of the background region. Fig. 10, however, has the following pattern[3]: white (W), medium crosshatching (MC'), heavy crosshatching (HC'), white (W), heavy crosshatching (HC'), and white (W). To make clear the relationship of Fig. 10 to Figs. 8-9, this order should be changed to W, LC', MC', W, MC', and W.

12. ·   The suggested changes to Figs. 9 and 10 would adequately resolve the issue of how the pattern in Fig. 10 is obtained and would, thereby, obviate any modification of the Specification to that effect – the pattern in Fig. 10 could be inferred from Figs. 8 and 9. Further support for these changes to Figs. 9 and 10 can be found in U.S. Patent 6,356,298, which discloses the coded structured light methodology utilized in the instant application. Specifically, the Applicant is referred to Figs. 7-10 of that patent.

Objections: Title of the Invention

13.   The title of the invention is not descriptive. A new title is required that is clearly indicative of the invention to which the claims are directed. The title should embody at least the following aspects of the Applicant's invention:

a.   The usage of coded structured light.

b.   The primary application of the invention to the visualization, extraction, and recognition of

handwritten characters.

c.   The capture of a time series of frames.

---

3   The ˙ notation denotes a different orientation of the crosshatching. For example, LC˙ indicates the same "degree" of crosshatching as LC –
i.e. light crosshatching – but the orientation of LC˙ crosshatching differs from LC (cf. the square region and background region of Fig. 8).
The intensity of the stripes may appear differently on the face of the box than on the wall. The Applicant likely intended to express this
difference through the orientation of the hatching that appears in the aforesaid square region and that which appears in the background
region. It is, therefore, understood that orientation of the crosshatching in the box region need not be the same as that of the background.

Claims

Response to Amendments to the Claims

14.   The amendments to the Specification, filed on the 7th of September 2004 (hereinafter, *Amendments to the*

Claims) have been entered and made of record. Claims 1, 3-4, 6-7, 9, 11-12, 14-15, 17, 21-22, 27, 32, and 37-40

have been amended accordingly. Claims 5 and 13 have been cancelled.

15.   As a result of these amendments, all prior claim objections have been withdrawn. Likewise, all prior 35

U.S.C. § 112(2) rejections have been withdrawn.

16.   Essentially, the Applicant has amended Claims 1 and 9 so as to incorporate the following limitations from

their respective dependent claims:

       a.   Comparison between frame data images picked up in a time-series (cf. original Claims 3-4

           and 11-12).

       b.   Storing, as initial frame data, an initial image of frame data picked up in a time-series

           transformed by the geometric transformation part (cf. original Claims 5 and 13).

       c.   Extracting only differential data as storage data based on a result of the comparison in the

           frame data comparison step between the initial frame data and frame data subsequently

           transformed in the time-series (cf. original Claims 5 and 13).

Claims 17, 27 and 37-40 have been similarly amended. Claims 17 and 27 have been amended to also include the

following limitation:

       d.   ...wherein the stored geometric-transformed intensity image is the initial geometric-

           transformed intensity image and the differential data between successive geometric-

           transformed intensity images in the time-series.

17.   Items (a)-(c) were addressed in the previous Office Action. Item (d) will be addressed below. The pending

claims are, therefore, rejected under 35 U.S.C. § 103(a) in view of the Prior Art cited in the previous Office Action[4].

Response to Arguments and Remarks

18.   The Applicant's arguments, filed on the 7th of September 2004 (hereinafter, *Applicant's Arguments*), have

been fully considered but they are not persuasive for at least the following reasons.

19.   The general thrust of the Applicant's arguments is that: "none of the recited prior art disclose or suggest an

image processing part that retrieves only differential data between successive geometric-transformed intensity

images [or frame data] in the time-series ... based on the result of the comparison of ... the initial geometric-

transformed intensity images [or initial frame data] and geometric-transformed intensity images [or frame data]

subsequently transformed in the time series"; nor do the recited prior art disclose or suggest that "the stored

geometric intensity image is the initial geometric-transformed image and the differential data between successive

geometric-transformed intensity images in the time series". These arguments, or arguments to this effect, are

reiterated throughout the Applicant's arguments. The Applicant provides brief synopses for each of the reference

used in the previous Prior Art rejections, with particular attention to U.S. Patent 5,764,383 (P. Perona et al., *U.S.*

Patent 5,764,383: Apparatus and Method for Tracking Handwriting from Visual Input – hereinafter, [Perona00]).

20.   In what follows, a thorough rebuttal of these arguments is presented. This is then followed by a brief

treatment of [Perona00] and its pertinence to the Applicant's claimed invention.

21.   As shown in the previous Office Action (cf. page 6, paragraph 13), [Kang95] (S.B. Kang, J.A. Webb, and

T. Kanade, *A MultiBaseline Stereo System with Active Illumination and Real-Time Image Acquisition*. IEEE, 1995)

discloses a method of depth reconstruction which utilizes a coded structured light (active illumination) to derive

stereo correspondences between images obtained from a multitude of cameras. It was argued that [Kang95], in

essence, constitutes an image processing method which satisfies the limitations of Claims 1 and 9. [Kang95],

however, does not show "retrieving only differential data between successive geometric-transformed intensity

---

[4]   [Bunke99] and [Mack00] have been introduced to address the newly added claim language.

images [or frame data] in the time-series ... based on the result of the comparison of ... the initial geometric-transformed intensity images [or initial frame data] and geometric-transformed intensity images [or frame data] subsequently transformed in the time series".

22.   As shown in the previous Office Action (cf. page 13, paragraphs 34-45), [Perona00] teaches "extracting only differential data ... based on the result of the comparison between the initial frame data and frame data subsequently picked up". For instance, [Perona00] suggests evaluating the difference, $F10 - F1$, between a subsequent frame $F10$ and the initial frame $F1$ ([Perona00] Fig. 4 steps **400-402**). The difference provides a measure of the *velocity* between frames. The velocity is concept related to the *optical flow* exhibited in temporal sequences of images, and essentially measures the amount of change observed in the image sequence over a specified duration. In the example above, $F10 - F1$ is indicative of the degree to which initial frame $F1$ changes over the duration spanning frames $F10$ and $F1$. According to [Perona00], the position exhibiting a maximal velocity (i.e. where the difference image $F10 - F1$ greatest magnitude or intensity) typically indicates the presence of the hand and writing implement. Please refer to [Perona00] column 5, lines 30-41. A neighborhood can be formed about the position of maximal velocity, whereby the actual position of the writing implement can be determined for the given frame, then predicted (e.g. using a Kalman Filter) and tracked (e.g. using correlative techniques) through successive frames.

23.   It was argued in the previous Office Action (page 13, paragraph 36), that the methods of [Perona00] and [Kang95] could be combined so as to track the three-dimensional position of the writing implement using the stereo-based technique of [Kang95]. [Perona00] suggests the usage of such techniques for precisely this purpose in the section entitled *Three-Dimensional Tracking*, and, more specifically, in column 8, lines 27-29. With the three-dimensional position of the writing implement in hand, one can determine whether or not it has been lifted off of the writing surface.

24.   Recall that, in order to facilitate the derivation of stereo correspondences between the captured stereo images, [Kang95] determine and administer a planar homography which transforms pairs of stereo images such that the resulting epipolar lines are parallel and equal along all scanlines of the transformed images. This process is referred to as *rectification* and is illustrated in Fig. 3 of [Kang95]. It was argued in the previous Office Action (cf. page 6, paragraph 13, item (9.c.)) that this process of rectification represented a type of geometric transformation.

25.   Notice that, in [Perona00], the determination of whether the writing implement is in an up or down position ([Perona00] Fig. 2, step 240) follows the determination of the its position within the given frame ([Perona00] Fig. 2; steps 210-230). Assuming [Kang95] is utilized in step 240, one could then conclude that the geometric transformation of the frames (e.g. frames $F1$ ... $F10$) follows the position determination of steps 210-230, and, hence, the evaluation of the image difference. In other words, frame data (of the time-series of frame data) is subsequently (i.e. subsequent to the retrieval of differential data) transformed. Taking into account the discussion above with respect to [Perona00], the combination of [Kang95] and [Perona00] thus involves "retrieving only differential data between successive frame data in the time-series [(e.g. $F10 - F1$)] ... based on a result of the ... comparison of the initial frame data [(e.g. $F1$)] and frame data [(e.g. $F10$)] subsequently transformed in the time series" (all images would be subsequently transformed, as discussed above). This is in accordance with the present language of Claim 1.

26.   The following is a brief observation as to the language of Claim 1. The Applicant should note the differences between the language of Claim 1 and that of Claim 9. Aside from the features of Claim 9 concerning "range information" – information, which is almost implied by the various imaging components of Claim 1 – the primary substantive difference between Claims 1 and 9 is that Claim 9 refers to "geometric-transformed intensity images", whereas, Claim 1 refers simply to "frame data". Clearly, intensity images can be interpreted as frame data, and vice versa. However, by distinguishing the intensity images as being "geometric-transformed", the language of Claim 9 implies that the intensity images have been geometric-transformed at some previous stage of the claimed image processing method. Therefore, according to Claim 9, the geometric transformation must precede the storing step, the comparison step, and the retrieval of the differential data.

27.   The language of Claim 1, on the other hand, does not necessitate this sequence. This was illustrated above with respect to the limitation, "retrieves only differential data ... frame data subsequently transformed ...". To illustrate this further, consider the limitation "... stores, as initial data, an initial image of frame data in a time series transformed by the geometric transformation part". This is not necessarily the same as "... stores, as initial data, an initial image of frame data in a time series *that have been* transformed by the geometric transformation part". Indeed, given its broadest, reasonable interpretation, the given limitation could be interpreted as "... stores, as initial data, an initial image of frame data in a time series *to be* transformed by the geometric transformation part". No

language of Claim 1 prohibits such an interpretation. With respect to Claim 1, this latter interpretation is adopted. Consequently, the rejections of Claim 1 and its dependant claims will follow in a similar manner as presented in the previous Office Action. However, with respect to Claim 9, this interpretation is clearly improper, since, as discussed above, the geometric transformation must precede the storing step, the comparison step, and the retrieval of the differential data. New grounds for rejection are present to address these limitations. In view of these new grounds of rejection the Applicant's remarks, with respect to Claim 9, 17, 27, and 37-40, are rendered moot.

28.   *[Perona00].* Briefly, [Perona00] discloses a system (and method) for detecting the movement of a writing implement relative to a writing surface to determine the path of the writing implement. The path of the writing implement can be used to define the handwriting of the user, which may be later recognized using OCR or other similar techniques. [Perona00] is distinguished over other systems in that a video camera is used to capture a sequence of images of the writing surface (cf. [Perona00] Fig. 4, step **400** and *Detailed Description of the Preferred Embodiments,* paragraph 2). These images can be images of a pen or other writing instrument, including the hand and/or fingers while it is *tracing letters*, graphic *characters*, or any other image formed by user's hand movement. Clearly (and notwithstanding the statement "this system preferably monitors relative movement of the writing implement, instead of imaging previously-written characters" – [Perona00] column 3, lines 43-46[5]), by capturing video images of the writing surface as the user writes, the captured sequence of images presumably would contain images depicting previously written characters. As discussed above, differential data (e.g. $F10 - F1$) is obtained between successive frames of the temporal sequence of frames. [Perona00] will be discussed further in the rejections that follow.

Rejections Under 35 U.S.C. § 103(a)

---

5   The Applicant has chosen, in his remarks, to emphasize this statement. This statement must be evaluated within the context of [Perona00], as a whole. What this statement means is that, instead of analyzing the image data corresponding to previously-written characters, or pieces thereof, [Perona00] chooses to capture the user's handwriting by monitoring the trajectory of the writing implement about the writing surface. As stated above, the images obtained by the video camera can include written characters, but they must include some depiction of the writing implement. Finally, note that the fact that [Perona00] monitors the relative motion of the writing implement, as opposed to the actual written characters, has little bearing on the reference's applicability to the Applicant's invention, as it is currently claimed.

29.   The following is a quotation of 35 U.S.C. § 103(a) which forms the basis for all obviousness rejections set

forth in this Office action:

(a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if
the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would
have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains.
Patentability shall not be negatived by the manner in which the invention was made.

30.   Claims 1, 4, and 8 are rejected under 35 U.S.C. § 103(a) as being unpatentable over [Kang95] in view of

[Perona00].

31.   *The following is in regard to Claim 1.* [Kang95] disclose an depth reconstruction system utilizing multiple

cameras and the projection of temporally coded structured light ([Kang95] Abstract). The system includes the

following:

(1.a.)   A projecting part that projects a pattern (i.e. a projector which casts a pattern of

sinusoidally varying intensity – [Kang95] Section 2, paragraph 1, last two sentences).

(1.b.)   Four cameras (i.e. a first, second, third, and fourth image pickup part) picks up stereo

intensity images of a scene illuminated by an active illumination (i.e. a "projection

pattern" of light having sinusoidally varying intensity). The cameras have disjoint optical

axes that are separated by some fixed *baseline* distance. See paragraph 1 of Section 2 on

page 88 of [Kang95]

(1.c.)   Creating first range information based on the pattern picked up by the second image

pickup part (i.e. recovering depth (range) data from analysis of the stereo images – cf.

Abstract of [Kang95]).

32.   Items (1.a.)-(1.c.) collectively qualify as a "three-dimensional image pickup part". The system further

includes:

(1.d.)   A geometric transformation part that performs a geometric transformation (i.e.

rectification – cf. [Kang95], Fig. 5 and Section 4.3).

33.   [Kang95] does not, however, disclose:

(1.e.)   A storage part that stores, as initial frame data, an initial image of frame data in a time-

series transformed by geometric transformation part.

(1.f.)   A frame data comparison part that makes a comparison between successive frame data

images in a time-series transformed by the geometric transformation part

(1.g.)   An image processing part that retrieves only differential data between successive frame

data in the time series as storage data based on a result of the comparison of the frame

data comparison part of the initial frame data subsequently transformed in the time-

series.

34.   [Perona00] disclose a method and apparatus for tracking handwriting by detecting the movement of a

writing implement relative to a writing surface relative to a writing surface. To capture the movements of the

writing, [Perona00] perform the following:

(1.e$_{Perona}$.)   A storage part that stores, as initial frame data, an initial image of frame data (i.e. frame

$F1$ – cf. [Perona00] Fig. 4, steps **400-402**) in a time-series (i.e. sequence of frames).

(1.f$_{Perona}$.)   A frame data comparison part that makes a comparison (e.g. evaluating the magnitude of

the difference between frames – i.e. the *velocity*; cf. [Perona00] column 5, lines 37-48)

between successive frame data images (e.g. frames $F1$ ...$F10$ ...) in a time-series.

(1.g$_{Perona}$.)   An image processing part that retrieves only differential data between successive frame

data in the time series as storage data based on a result of the comparison of the frame

data comparison part of the initial frame data (cf. [Perona00] Fig. 4, steps **400-402** and

column 5, lines 30-41).

[Perona00] also suggests a stereo-vision approach to determining the 3D position (particularly depth information) of

the writing implement during the tracking ([Perona00] column 7, lines 66-67 to column 8, lines 1-6 and 23-28).

35.   The teachings of [Perona00] and [Kang95] are combinable because they are analogous art. [Perona00]

suggest the usage of stereo-vision techniques to acquire 3D information of the writing surface, while [Kang95]

acquiring 3D information of an observed scene via stereo-vision techniques. Therefore, it would have been obvious

to one of ordinary skill in the art, at the time of the Applicant's claimed invention, to extend the stereo-vision depth

recovery methodology of [Kang95] to the handwriting tracking method of [Perona00] The motivation to do so

would have been to exploit the accuracy of [Kang95] method of depth recovery ([Kang95] Abstract) to detect and

track the 3D dimensional position of a writing implement, according to [Perona00]'s method.

36.   As discussed in the previous section of this document, frames would be subsequently transformed in this

combination. Following that from discussion, it can be concluded the combination of [Kang95] and [Perona00]

would include:

(1.e.)   A storage part that stores, as initial frame data, an initial image of frame data in a time-

series transformed by geometric transformation part.

(1.f.)   A frame data comparison part that makes a comparison between successive frame data

images in a time-series transformed by the geometric transformation part

(1.g.)   An image processing part that retrieves only differential data between successive frame

data in the time series as storage data based on a result of the comparison of the frame

data comparison part of the initial frame data subsequently transformed in the time-

series.

Recall the manner in which Claim 1 is being interpreted. See the discussion in the previous section of this document.

37.   *The following is in regard to Claim 4.* The process of *rectification* (cf. [Kang95] page 89, Section 4,

paragraphs 1-2 and Fig. 3) involves modifying a position of the frame data image based on a result of the

comparison between the frame data images.

38.   *The following is in regard to Claim 8.* The cameras used in the method of [Kang95] have non-parallel

optical axes. See, for example, [Kang95] Fig. 3 and the description of the *verged camera configuration* in Section

2.1. Therefore, the cameras (including the second camera or image pickup part) capture the measurement target at

different angles. Depth (range) information, according to the method of [Kang95], is derived from the images of the

scene having active illumination projected onto it. Taking this into account, the method of [Kang95] is such that the

cameras (the image pickup parts), including the second camera, comprise plural image pickup parts that pick up the

measurement target at different angles. The method also involves creating range information based on projection

patterns respectively picked up by the plural image pickup parts of the cameras, including the second camera.

39.   Claims 2-3 and 7 are rejected under 35 U.S.C. § 103(a) as being unpatentable over [Kang95] and

[Perona00], as applied to Claim 1, in further view of [Batlle98] (J. Battle et al., *Recent Progress in Coded Structured*

Light as a Technique to Solve the Correspondence Problem: A Survey. Pattern Recognition, Vol. 31, No. 7, pp. 963-

982, 1998).

40.   *The following is in regard to Claim 2.* As shown above, [Kang95] disclose a depth recovery system that

satisfies the limitations of claim 1. [Kang95], however, does not expressly show or suggest that the range

information be derived by:

   (2.a.)   For an area where the amount of change of the pattern picked up by the first image

               pickup part with respect to the projection pattern is equal to or greater than a

               predetermined value.

   (2.b.)   Assigning new code corresponding to the pattern picked up by the first image pickup

               part.

   (2.c.)   Creating the first range information from the pattern picked up by the second image

               pickup part based on the new code.

41.   [Batlle98] present a survey of various coded structured light methods. With regard to claim 2, reference

will be made here to the technique suggested by Posdamer and Altshuler and discussed in Section 6.1 (page 969) of

[Batlle98], as well as Sato and Inokuchi's (Sato et al.) later extension of that methodology ([Batlle98] page 970).

The technique includes the projection of a temporally varying binary pattern onto the scene ([Batlle98] Fig. 3). In

this way, sections of the patterns are temporally encoded ([Batlle98] Fig. 4). Sato et al. propose the usage of a

dynamic threshold indicative of the difference between the observed pattern intensity and the reference (projected)

pattern intensity ([Batlle98] page 970, last paragraph of left column). Taking this into account, [Batlle98] thus

teaches a coded structured light method that includes:

   (2.a$_{Batlle}$.)   For an area where the amount of change of the pattern picked up by the an image pickup

               part with respect to the projection pattern is equal to or greater than a predetermined

value ([Batlle98] page 970, last paragraph of left column).

(2.b$_{Batlle}$.)   Assigning new code corresponding to the pattern picked up by the image pickup part

(e.g. by the temporal binary codification discussed above).

(2.c$_{Batlle}$.)   Creating the first range information from the pattern picked up by the image pickup part

based on the new code ([Batlle98] Abstract).

42.   The teachings of [Batlle98] and [Kang95] are combinable because they are analogous art. Specifically,

[Kang95] describe a method that uses structured light principles in a passive stereo system, yet omit the details of

the structured light employed. [Batlle98], on the other hand, discuss in detail various structured light methods,

including one that involves steps (2.a$_{Batlle}$.)- (2.c$_{Batlle}$.) above. Therefore, it would have been obvious to one of

ordinary skill in the art, at the time of the Applicant's claimed invention, to utilize the coded structured light

methodology discussed above and by [Batlle98] to codify the projected pattern used in [Kang95]. The motivation to

do so would have been to distinguish pixels or groups of pixels in the captured images by determining which beam

column (or other coded pattern segment) has been projected onto the scene ([Batlle98] Section 6.1, paragraph 3 (last

sentence)). This increases the local discriminability at each pixel or pixel group, thereby, allowing the acquisition of

accurate and dense depth data ([Kang95] page 88, right column, sentence 1). Using such a coded structured light

scheme in the depth recovery method of [Kang95] yields a method where range data is derived by:

(2.a.)   For an area where the amount of change of the pattern picked up by the four cameras

(including the first camera) with respect to the projection pattern is equal to or greater

than a predetermined value.

(2.b.)   Assigning new code corresponding to the pattern picked up by the cameras (including the

first camera).

(2.c.)   Creating the first range information from the pattern picked up by the cameras (including

the second camera) based on the new code.

43.   *The following is in regard to Claim 3.* According to [Batlle98] (cf. [Batlle98] page 978, right column, lines

1-3), the presence of noise obfuscates the derivation of the correspondence between the images captured from the

cameras. Therefore, in order to facilitate the derivation of a proper correspondence and, thereby, eliminate potential

depth discontinuities, it would have been obvious to one of ordinary skill in the art, at the time of the applicant's

claimed invention, to eliminate noise data from the frame data image based on a result of the comparison between

the frame data images.

44.   *The following is in regard to Claim 7.* In the method of [Kang95], as in most methods utilizing stereo

vision, correspondences are derived between captured stereo images, in processes known as *stereo matching.* See,

for example, [Kang95] Abstract, Introduction (paragraph 1), and paragraph 1 of Section 4.1, in conjunction with

Figs. 3-4. Note also that such correspondences are also embodied in the various homographies derived according to

the method of [Kang95] (paragraph 3 of [Kang95] Section 4). These homographies provide a bijective map between

points, and hence *all* areas, of adjacent images. These homographies are derived between all adjacent cameras

(including the first and second). [Kang95], however, do not show determining an amount of change of the pattern

captured by a camera with respect to the projection pattern and determining whether or not it is less than a

predetermined value.

45.   As discussed above (see the discussion above with respect to steps (2.a)-(2.c)), [Batlle98] teaches a coded

structured light method that includes: Evaluating whether an area is such that the amount of change of the pattern

picked up by the an image pickup part with respect to the projection pattern is equal to or greater than a

predetermined value ([Batlle98] page 970, last paragraph of left column). Consequently, the determination of

whether not the amount of change is less than the predetermined value is made implicitly for areas deemed not to

satisfy this criterion.

46.   It would have been obvious to one of ordinary skill in the art, at the time of the Applicant's claimed

invention, to evaluate whether or not an area is such that the amount of change of the pattern picked up by the an

image pickup part with respect to the projection pattern is less than a predetermined value. As indicated in [Batlle98]

([Batlle98] page 970, last paragraph of left column), this has the advantage of distinguishing between high contrast

transitions that are part of the observed scene itself, as opposed to high contrast transitions that occur as a result of

the projected pattern. Combining the teachings of [Batlle98] and [Kang95] yields a method of depth recovery that

includes:

(7.a.)   Evaluating whether or not areas of a captured images is such that the amount of change

of the pattern picked up by the an image pickup part with respect to the projection

pattern is less than a predetermined value.

(7.b.)   Deriving correspondences between *all* adjacent (including those obtained from the first

and the second cameras) stereo images (including those obtained from the first and the

second cameras).

Since, in such a method, correspondences are derived for all pixels in the captured images (including those having

an amount of change less than the predetermined value), the system obtained as such conforms sufficiently to that of

Claim 7.

47.   Claims 6 is rejected under 35 U.S.C. § 103(a) as being unpatentable over [Kang95] and [Perona00], as

applied to Claim 1, in further view of [Mack00] (W. Mack et al., *U.S. Patent 6,125,197: Method And Apparatus For*

The Processing Of Stereoscopic Electronic Images Into Three-Dimensional Computer Models Of Real-Life Objects.

Filing Date: June 1998).

48.   *The following is in regard to Claim 6.* As shown above, [Kang95] discloses a depth recovery system that

satisfies the limitations of Claim 1. Though [Kang95] shows the capture of the intensity image of the scene having

an active illumination projected onto it (i.e. the pattern projection image and intensity image picked up in parallel –

see above), [Kang95] does not expressly show or suggest that:

(6.a.)   The projecting part has a light source emitting light of an invisible region of a

wavelength band.

(6.b.)   The first and second image pickup parts have a filter for transmitting light of the invisible

region of the wavelength band and a filter for cutting off light of an invisible region of

the wavelength band, and pick up the projection pattern image and intensity image in

parallel.

49.   [Mack00] discloses a structured light method and apparatus for extracting three-dimensional (3-D) data

from a scene (cf. [Mack00], Abstract and Fig. 4). [Mack00] suggests projecting structured light patterns in the infra-

red region of the spectrum (i.e. "the projecting part has a light source emitting light of an invisible region of a

wavelength band"). Furthermore, in order for the multiple stereo cameras (e.g. cameras 22, 23, and 24 of [Mack00]

Fig. 2) to simultaneously capture both visible and infra-red light, [Mack00] suggests using an appropriate set of

color filters ([Mack00] column 5, lines 25-30), or a composite filter, wherein one portion of the filter bars the

transmission of visible-spectrum radiation, while the another portion bars transmission of IR-spectrum radiation (cf.

[Mack00] column 5, lines 39-57).

50.   The teachings of [Mack00], [Perona00], and [Kang95] are clearly combinable, in the sense that each

discloses or suggests some method or application of stereo image processing. In particular, [Mack00] and [Kang95]

disclose stereo reconstruction methods involving structured light. Therefore, it would have been obvious to one of

ordinary skill in the art, at the time of the Applicant's claimed invention, to project structured light patterns

consisting of infra-red radiation, and to further affix composite filters, such as those employed in [Mack00], to each

of the multiple stereo cameras. The advantage of such a configuration is twofold. First, infra-red radiation is not

visible to the naked eye. Infra-red light patterns, therefore, would not distract a user as visible light patterns clearly

could. This would be particularly advantageous in handwriting applications such as [Perona00]. (Infrared imagers

are also less susceptible to artifacts due to specular reflections, shadows, and poor ambient lighting conditions).

Secondly, if infra-red patterns were to be used, the usage of composite filters, like the ones described above, would

obviate the need for both visible light cameras and infra-red cameras. The resulting reduction in hardware reduces

the overall complexity of the system.

51.   Claims 9, 11-12, 14, and 16 are rejected under 35 U.S.C. § 103(a) as being unpatentable over [Bunke99]

(H. Bunke et al., *Online Handwriting Data Acquisition Using a Video Camera*, Fifth International Conference on

Document Analysis and Recognition: ICDAR '99, September 1999), in view of [Saund98(383)] (E. Saund et al.,

U.S. Patent 5,764,3873: Platenless Book Scanner With Line Buffering To Compensate For Image Skew, Filing Date:

May 1996)[6], and in further view of [Mack00].

52.   *The following is in regard to Claim 9.* [Bunke99] introduces a method for online data acquisition of

handwriting, wherein a video camera is used to record the handwriting of a user. Essentially, [Bunke99] is an image

processing method. The method of [Bunke99] includes:

(9.a$_{Bunke}$.)   Storing an initial intensity image in a time-series. A temporal sequence of images is

captured using the video camera (cf. [Bunke99] Section 2, *System Description*, paragraph

1). These images, like all images, can be considered "intensity images", and a sequence

of such images would inherently contain an "initial intensity image".

The *differential image processing* stage of [Bunke99] involves the following:

(9.b$_{Bunke}$.)   Making comparisons between successive intensity images (i.e. consecutive images taken

at, say, time t and at time $t+1$ – cf. [Bunke99] Section 2.1, *Differential Image Processing*,

paragraph 1) in the time-series. These comparisons include, for example, evaluating the

differences between the two consecutive images (cf. [Bunke99] Section 2.1, *Differential*

Image Processing, paragraphs 1-2)

(9.c$_{Bunke}$.)   Retrieving only differential data (i.e. *differential images* – e.g. [Bunke99] Fig. 1) between

successive intensity images in the time series as storage data based on a comparison of

the comparison step of the initial intensity image and the subsequent intensity images of

the time-series. The differential images (i.e. the difference between two images) obtained

over a certain duration (e.g. from time t to $t+1$) provide a representation of the trace of

ink over that span of time. In this sense, *only* the differential images are necessary to

construct a complete trace of user's handwriting (cf. [Bunke99] Section 2.1, *Differential*

---

6   Notice that, in column 8, lines 31-37 of [Saund98(383)], U.S. Patent Application 09/657,711 (now U.S. Patent 5,760,925 issued to the same inventors and assignee) is incorporated by reference. In accordance with MPEP § 2163.07(b), all material disclosed therein will be treated as part of U.S. Patent 5,764,383. Consequently, when referring to [Saund98(383)] in this document, reference will actually be made to both U.S. Patent 5,764,383 and U.S. Patent 5,760,925. For the sake of clarity, any specific references to either U.S. Patent 5,764,383 or U.S. Patent 5,760,925 will be denoted as [Saund98(383)] or [Saund98(925)], respectively.

Image Processing, paragraph 1).

[Bunke99] does not, however, show or suggest:

(9.d.)   Projecting a pattern by a projecting part.

(9.e.)   Picking up an intensity image and a projection pattern image by a first image pickup part

from an optical axis direction of the projecting part, and picking up the projection pattern

image by a second image pickup part from a direction different from the optical axis

direction of the optical axis direction.

(9.f.)   Creating first range information based on the pattern picked up by the second image

pickup part.

(9.g.)   Performing a geometric transformation for the intensity image produced by the first

image pickup part based on the range information.

53.   [Saund98(383)] discloses a system for scanning bound documents, face-up and in an open condition (cf.

[Saund98(383)] Fig. 1). In particular, the scanning system of [Saund98(383)] is capable of de-warping skewed

images of the bound documents (cf. [Saund98(383)], *Summary of the Invention*, paragraph 1). In order to

accomplish this correction, [Saund98(383)] requires the three-dimensional (3D) shape of the page. The 3D shape of

the page is derived using a structured light technique ([Saund98(383)], *Detailed Description*, paragraph 2), wherein

a pattern of at least two stripes is projected across the page and an image thereof (i.e. *geometric page shape data* –

[Saund98(383)] column 5, lines 65-67 and column 6, lines 9-12) is obtained by the camera 15. An image of the page

without the projected pattern (i.e. *raw image data* – [Saund98(383)] column 5, line 67) is also captured. Therefore,

[Saund98(383)] performs the following:

(9.d.)   Projecting a pattern by a projecting part.

(9.e$_{Saund}$.) . Picking up an intensity image (i.e. raw image data) and a projection pattern image (i.e.

geometric page shape data) by a single image pickup part (camera 15).

Using the geometric page shape data, [Saund98(383)] determines the 3D shape of the document. The 3D shape is

expressed essentially in terms of a "geometric transform" called the *page shape transform* (cf. [Saund98(383)]

column 6, lines 11-14). Once the 3D shape is known, the amount of skew (cf. [Saund98(383)] Figs. 5-6) is measured

(in a 3D *world space* – [Saund98(383)] column 8, lines 1-3), and the raw image can be corrected in accordance with

the determined skew, the page shape transform determined by calculator **34**, and the perspective transform **42**

([Saund98(383)] column 6, lines 45-47). The page shape transform, the perspective transform, and the skew

correction all represent "geometric transformations". [Saund98(925)] discusses these processes in greater detail and

suggests methods for dealing with more general deformations such as foreshortening or magnification due to the

curved shape of the document pages (cf. [Saund98(383)], column 8, lines 48-53 and [Saund98(925)] Abstract).

Corrections of this nature also represent geometric transformations. Thus, [Saund98(383)] clearly shows:

(9.f$_{Saund}$.)   Creating first range information based on the pattern picked up by the image pickup part.

(9.g$_{Saund}$.)   Performing a geometric transformation (e.g. correction of skew, correction of

foreshortening, or correction of magnification) for the intensity image produced by the

image pickup part (e.g. camera **15**) based on the range information (i.e. the 3D shape of

the bound document).

54.   [Bunke99] and [Saund98(383)] are combinable because they are analogous art. In particular, both

[Bunke99] and [Saund98(383)] disclose image-based methods for observing and analyzing writing surfaces

containing or capable of accepting text. For example, the bound document in [Saund98(383)] could be a writing

surface such as a ledger or notebook. Therefore, it would have been obvious to one of ordinary skill in the art, at the

time of the Applicant's claimed invention, to modify [Bunke99] by using the scanning method of [Saund98(383)] to

compensate for distortions – such as skew, magnification and/or foreshortening – in the individual frames of the

aforesaid video sequence. The method of [Saund98(383)] is an alternative to the morphological operations

[Bunke99] uses to compensate for "small displacements" ([Bunke99], Section 2.1, paragraph 3, last sentence).

Moreover, by using the scanning method of [Saund98(383)], the method of [Bunke99] would no longer be restricted

to analyzing writing surfaces which are flat (essentially two-dimensional). Instead, the user could write on bound

documents such as those shown in Figs. 1 and 3. Bound documents, such as notepads and ledgers, are a common

form of handwriting media.

55.   One of ordinary skill in the art would realize that, by incorporating [Saund98(383)] into [Bunke99] in this

manner, the elimination of distortions – and, hence, the application of the various geometric transformations – must

precede the storage of the individual image frames (i.e. step (9.a.) or step (9.a$_{Bunke}$.) above). Otherwise, these

distortions could propagate to the subsequent steps and corrupt the trace of the user's handwriting. Taking this into account, the combination of [Bunke99] and [Saund98(383)] would include the following:

(9.d.)   Projecting a pattern by a projecting part.

(9.e$_{Saund}$.)   Picking up an intensity image (i.e. raw image data) and a projection pattern image (i.e. geometric page shape data) by a single image pickup part (camera 15).

(9.f$_{Saund}$.)   Creating first range information based on the pattern picked up by the image pickup part.

(9.g$_{Saund}$.)   Performing a geometric transformation for the intensity image produced by the image pickup part based on the range information.

(9.a.)   Storing an initial *geometrically-transformed* intensity image in a time-series.

(9.b.)   Making comparisons between successive *geometrically-transformed* intensity images in the time-series *transformed in the geometric transformation step.*

(9.c.)   Retrieving only differential data between successive *geometrically-transformed* intensity images in the time series as storage data based on a comparison of the comparison step of the initial *geometrically-transformed* intensity image and the *geometrically-transformed* intensity images *subsequently transformed* (i.e. transformed subsequent to the transformation of the initial intensity image) in the time-series.

Notice, however, that [Saund98(383)] employs only a single camera for capturing the projection pattern image and the intensity image, not first and second "image pickup parts", as claimed.

56.   [Mack00] discloses a structured light method and apparatus for extracting three-dimensional (3-D) data from a scene (cf. [Mack00], Abstract and Fig. 4). [Mack00] illustrates several configurations for 3D imaging devices that use structured light. For example, the configuration depicted in Fig. 2 consists of a projector (elements 16-18), and three cameras – cameras 22-24. The functionality of this configuration is discussed in column 6, paragraph 2. Accordingly, each camera – camera 22, camera 23, and camera 24 – is configured to capture the structured-light pattern projected onto the given scene ([Mack00], column 6, lines 31-34). One camera, camera 24, is configured to obtain a color image (i.e. an "intensity image") of the scene ([Mack00], column 5, lines 26-28). Notice that the optical axis of camera 24 is parallel to that of the projector. Also, notice the optical axes of the other cameras are

oriented in a different direction. Therefore, camera **24** can be treated as the "first image pickup part" and either of

the cameras **22** or **23** can be considered the "second image pickup part", in accordance with Claim 9.

57.   Such a configuration is preferable over the single camera configuration of [Saund98(383)] because images

of the structured light pattern and an image of the scene can be captured simultaneously ([Mack00] column 6, lines

31-34). Furthermore, by combining the information of all cameras, a 3D model of high resolution can be obtained

([Mack00] column 6, lines 42-45). This is a general result in stereo imaging and depth recovery – the more stereo

images available, or correspondingly the more stereo imagers available, the denser the map of recovered depth

values. Therefore, it would have been obvious to one of ordinary skill in the art, at the time of the Applicant's

claimed invention, to modify the combination of [Bunke99] and [Saund98(383)] so as to use a structured light

configuration similar to the one depicted in [Mack00] Fig. 2, when determining the 3D structure of the writing

surface or bound document. Recall that the goal of the structured light in [Saund98(383)] was to obtain the 3D shape

of the bound document. Any structured light method capable of this functionality would suffice and would not

substantially alter the remaining elements of [Saund98(383)] (or [Bunke99] , for that matter). The combination of

[Mack00], [Saund98(383)], and [Bunke99] would, therefore, include:

(9.e.)   Picking up an intensity image and a projection pattern image by a first image pickup part

. (e.g. [Mack00] camera **24**) from an optical axis direction of the projecting part, and

picking up the projection pattern image by a second image pickup part (e.g. [Mack00]

camera **23**) from a direction different from the optical axis direction of the optical axis

direction.

(9.f.)   Creating first range information based on the pattern picked up by the second image

pickup part (e.g. [Mack00] camera **23**).

(9.g.)   Performing a geometric transformation for the intensity image produced by the first

image pickup (e.g. [Mack00] camera **24**) part (e.g. [Mack00] camera **24**) based on the

range information.

58.   *The following is in regard to Claim 11:* [Bunke99] eliminate noise data (e.g. artifacts due to movements of

the hand, pen, or shadows induce by either – cf. [Bunke99] Section 2.1, *Problem 1* and *Problem 2*; and, on the same

page, the first and second paragraphs[7] in the right column) from the frame data image based on a result from the

comparison (e.g. evaluating the differences between consecutive images – cf. [Bunke99] Section 2.1, paragraphs 1-

2) between the frame data images in the frame data comparison step.

59.   *The following is in regard to Claim 12.* By correcting for skew, magnification, foreshortening, etc. (see

above), the position of the frame data image is modified based on the result of the comparison between the frame

images in the frame data comparison step.

60.   *The following is in regard to Claim 14.* As discussed above with respect to Claim 6, [Mack00]

demonstrates the usage of infra-red structured light patterns. [Mack00] also discloses capturing the intensity image

and pattern projection image simultaneously (i.e. "in parallel").

61.   *The following is in regard to Claim 16.* The configuration shown in Fig. 2 of [Mack00] includes:

(16.a.)   Plural (two) image pickup parts (i.e. camera **22** and camera **23**) that pick up the scene at

different angles.

(16.b.)   Creating range information based in projection patters respectively picked up by the plural

image pick up parts. Cameras **22** and **23** are configured to detect the projected pattern (cf.

[Mack00] column 6, lines 19-22). [Mack00] uses the images captured from these cameras to

derive 3D (range) information corresponding to the observed scene (cf. [Mack00] column 6,

lines 39-45).

Collectively, the camera **22** and camera **23** can be considered a "second image pickup part".

62.   Claims 10 and 15 are rejected under 35 U.S.C. § 103(a) as being unpatentable over [Bunke99],

[Saund98(383)], and [Mack00], as applied to Claim 9 above, in further view of [Batlle98].

---

7   When referring to paragraphs in the cited references, the convention followed here is that the paragraph number is assigned to paragraphs of
a given column (if applicable) or section, sequentially, beginning with the first full paragraph. Paragraphs that carry over to other columns
will be referred to as the last paragraph of the column in which they began.

63.   *The following is in regard to Claim 10.* Neither [Bunke99], [Saund98(383)], nor [Mack00] expressly show or suggest:

   (10.a.)   For an area where the amount of change of the pattern picked up by the first image

             pickup part with respect to the projection pattern is equal to or greater than a

             predetermined value.

   (10.b.)   Assigning new code corresponding to the pattern picked up by the first image pickup

             part.

   (10.c.)   Creating the first range information from the pattern picked up by the second image

             pickup part based on the new code.

64.   [Batlle98] present a survey of various coded structured light methods. With regard to claim 10, reference will be made here to the technique suggested by Posdamer and Altshuler and discussed in Section 6.1 (page 969) of [Batlle98] and Sato and Inokuchi's later extension of that methodology ([Batlle98] page 970). The technique includes the projection of a temporally varying binary pattern onto the scene ([Batlle98] Fig. 3). In this way, sections of the patterns are temporally encoded ([Batlle98] Fig. 4). Sato et al. propose the usage of a *dynamic threshold* indicative of the difference between the observed pattern intensity and the reference (projected) pattern intensity ([Batlle98] page 970, last paragraph of left column). Taking this into account, [Batlle98] thus teaches a coded structured light method that includes:

   (10.a$_{Batlle}$.)   For an area where the amount of change of the pattern picked up by the an image pickup

                 part with respect to the projection pattern is equal to or greater than a predetermined

                 value ([Batlle98] page 970, last paragraph of left column).

   (10.b$_{Batlle}$.)   Assigning new code corresponding to the pattern picked up by the image pickup part

                 (e.g. by the temporal binary codification discussed above).

   (10.c$_{Batlle}$.)   Creating the first range information from the pattern picked up by the image pickup part

                 based on the new code ([Batlle98] Abstract).

65.   The teachings of [Batlle98], [Bunke99], [Saund98(383)], and [Mack00] are combinable because they are analogous art. Specifically, [Batlle98] and [Saund98(383)] both relate to using structured light techniques to derive

3D information about an object or sceene. Therefore, it would have been obvious to one of ordinary skill in the art, at the time of the applicant's claimed invention, to utilize the coded structured light methodology discussed above and by [Batlle98], in lieu of the simplistic structured light method used in [Saund98(383)] to derive the 3D shape of the aforesaid writing surface or bound document. The motivation to do so would have been to distinguish pixels or groups of pixels in the captured images by determining which beam column (or other coded pattern segment) has been projected onto the scene ([Batlle98] Section 6.1, paragraph 3 (last sentence)). Generally speaking, coded structured light encodes more information into the projected patterns than un-coded structured light. Each discrete location of the pattern is associated with a code, which essentially consists of a unique sequence of colors, as opposed to a single color like in conventional structured light methods. This substantially improves the resolvability of correspondences between the projected pattern and the captured image depicting this pattern on the given scene. Using such a coded structured light scheme in the combination of [Bunke99], [Saund98(383)], and [Mack00] yields a method where range data is derived by

(10.a.)   For an area where the amount of change of the pattern picked up by any of the image

pickup parts (including the first image pickup part) with respect to the projection pattern

is equal to or greater than a predetermined value.

(10.b.)   Assigning new code corresponding to the pattern picked up by any of the image pickup

parts (including the first image pickup part).

(10.c.)   Creating the first range information from the pattern picked up by all of the image pickup

parts (including the second image pickup part) based on the new code.

66.   *The following is in regard to Claim 15.* All structured light methods attempt to solve what is known as the correspondence problem, which can be summarized briefly as: *for a pair of images, which regions of the images are projections of the same element of the three-dimensional scene?* See, for example, [Mack00] Fig. 12. Two images, 14 and 15, of the pattern projected scene 20 are captured by cameras 12 and 13, respectively. Note that these cameras could be cameras 22 and 23 of the 3D imaging configuration discussed above, as they image a pattern projected scene. In Fig. 12, the 3D point *D3* projects to points *D2* and *D1* in the image planes of cameras 12 and 13, respectively. Points *D2* and *D1* are referred to as stereo correspondences between the images 14 and 15. Having

knowledge of these correspondences and certain parameters of the camera configuration, one can determine the

point $D3$ by triangulation. By repeating this process for several points in the given images, one obtains range

information that delineates the observed scene. In other words, range information is created by deriving

correspondences between intensity information obtained by a first and second image pickup part. This process is

common to most stereo imaging and/or structured light methods. However, neither [Bunke99], [Saund98(383)], nor

[Mack00] expressly show or suggest:

    (15.a.)   For an area where the amount of change of the pattern picked up by the first image

        pickup part with respect to the projection pattern is less than a predetermined value,

        creating second range information by deriving a correspondence between intensity

        information obtained by the first and second image pickup parts.

67.   As discussed above (see the discussion above with respect to steps (10.a)-(10.c)), [Batlle98] teach

a coded structured light method that includes: Evaluating whether an area is such that the amount of change of the

pattern picked up by the an image pickup part with respect to the projection pattern is equal to or greater than a

predetermined value ([Batlle98] page 970, last paragraph of left column). Consequently, the determination of

whether not the amount of change is less than the predetermined value is made implicitly for areas deemed not to

satisfy this criterion. It would have been obvious to one of ordinary skill in the art, at the time of the applicant's

claimed invention, to evaluate whether or not an area is such that the amount of change of the pattern picked up by

the an image pickup part with respect to the projection pattern is less than a predetermined value. As indicated by

[Batlle98] ([Batlle98] page 970, last paragraph of left column), this has the advantage of distinguishing between

high contrast transitions that are part of the observed scene itself, as opposed to high contrast transitions that occur

as a result of the projected pattern. Combining the teachings of [Batlle98] with those of [Bunke99], [Saund98(383)],

and [Mack00] yields a method of depth recovery that includes:

    (15.a.)   Evaluating whether or not areas of a captured images is such that the amount of change

        of the pattern picked up by the an image pickup part with respect to the projection

        pattern is less than a predetermined value.

In creating range data of real objects, [Mack00] suggests that a multitude of images of real objects are taken from

different positions to exploit the differences in the objects' projection ([Mack00], column 2, lines 50-55). Since the

configuration above consists of three cameras, it would have been obvious to one of ordinary skill in the art, at the

time of the Applicant's claimed invention, to derive correspondences between the images obtained from each camera

(including those obtained from the first and the second cameras). As mentioned above, this is beneficial because the

more stereo images used, the denser the set of range information. In this case, correspondences are derived for all

pixels in the captured images, including areas where the amount of change of the pattern picked up by the first

image pickup part with respect to the projection pattern is less than a predetermined value.

68.   Claims 17-20, 27-30, and 37-40 are rejected under 35 U.S.C. § 103(a) as being unpatentable over

[Bunke99], in view of [Saund98(383)][8]

69.   *The following is in regard to Claims 17 and 27.* [Bunke99] disclose a method for online data acquisition of

handwriting, wherein a video camera is used to record the handwriting of a user. The method comprises:

(27.f_{Bunke}.)   Extracting a difference (i.e. differences between two consecutive frames) between an

intensity image (e.g. frame at time $t+1$) and an intensity image acquired in advance (e.g.

frame at time t) – cf. [Bunke99] Section 2.1 paragraphs

(27.g_{Bunke}.)   Storing, as the intensity image, an initial image in a time-series (cf. item (9.a_{Bunke}.)

above).

(27.h_{Bunke}.)   Making comparison between successive intensity images in the time-series (cf. item

(9.b_{Bunke}.) above).

(27.i_{Bunke}.)   Retrieving only differential data between successive intensity images in the time-series as

storage data (cf. item (9.c_{Bunke}.) above).

Furthermore,

(27.j_{Bunke}.)   The stored intensity image (i.e. *aggregate image* – [Bunke99] page 2, right column,

---

8   Notice that, in column 8, lines 31-37 of [Saund98(383)], U.S. Patent Application 09/657,711 (now U.S. Patent 5,760,925 issued to the same
inventors and assignee) is incorporated by reference. In accordance with MPEP § 2163.07(b), all material disclosed therein will be treated as
part of U.S. Patent 5,764,383. Consequently, when referring to [Saund98(383)] in this document, reference will actually be made to both
U.S. Patent 5,764,383 and U.S. Patent 5,760,925. For the sake of clarity, any specific references to either U.S. Patent 5,764,383 or U.S.
Patent 5,760,925 will be denoted as [Saund98(383)] or [Saund98(925)], respectively.

second to last paragraph, last sentence) is the initial intensity image and the differential

data between successive images in the time-series.

[Bunke99], however, does not show or suggest:

(27.a.)   Projecting light to an image holding medium to form an image thereon.

(27.b.)   Capturing the image projected on the image holding medium.

(27.c.)   Acquiring an intensity image based on the image picked up in the image pickup step.

(27.d.)   Acquiring range information from the picked-up image .

(27.e.)   Performing geometric transformation for the intensity image based on the range

information acquired in the range information acquisition step.

70.   [Saund98(383)] disclose a platenless book scanning system and method that uses structured light to obtain

a 3D representation of the observed book (cf. [Saund98(925)] Abstract). The method of [Saund98(383)] includes the

following:

(27.a.)   Projecting light to an image holding medium (e.g. platform 8 or bound document 10 of

[Saund98(383)] or [Saund98(925)] Fig. 1) to form an image thereon. See, for example,

[Saund98(925)] column 6, lines 19-35.

(27.b.)   Capturing the image (e.g. image I_1 – [Saund98(925)] column 7, lines 49-50) projected on

the image holding medium. See, for example, [Saund98(925)] column 6, lines 6-18.

(27.c.)   Acquiring an intensity image (e.g. image I_2 – [Saund98(925)] column 7, lines 53-54)

based on the image picked up in the image pickup step. See, for example,

[Saund98(925)] column 6, lines 6-18.

(27.d.)   Acquiring range information from the picked-up image (cf. [Saund98(925)] column 6

(lines 19-21), column 7 (lines 16-21) and Figs. 8-9).

(27.e.)   Performing geometric transformation (e.g. skew correction via the *page shape transform*)

for the intensity image based on the range information acquired in the range information

acquisition step. Refer to [Saund98(383)] Fig. 7 and column 7, lines 63-67 to column 8,

lines 1-8.

71.   [Saund98(383)] and [Bunke99] are combinable for the same reasons stated above. It would have been

obvious to one of ordinary skill in the art, at the time of the Applicant's claimed invention, to modify [Bunke99] by

using the scanning method of [Saund98(383)] to compensate for distortions – such as skew, magnification and/or

foreshortening – in the individual frames of the aforesaid video sequence. The method of [Saund98(383)] is an

alternative to the morphological operations [Bunke99] uses to compensate for "small displacements" ([Bunke99],

Section 2.1, paragraph 3, last sentence). Moreover, by using the scanning method of [Saund98(383)], the method of

[Bunke99] would no longer be restricted to analyzing writing surfaces which are flat (essentially two-dimensional).

Instead, the user could write on bound documents such as those shown in Figs. 1 and 3. Bound documents, such as

notepads and ledgers, are a common form of handwriting media.

72.   For the same reasons given above, the geometric transformation step (27.e.) would proceed steps

(27.f_{Bunke}.)-(27.j_{Bunke}.) above. Taking this into account, the combination of [Bunke99] and [Saund98(383)] would

include the following:

       (27.a.)   Projecting light to an image holding medium to form an image thereon.

       (27.b.)   Capturing the image projected on the image holding medium. ·

       (27.c.)   Acquiring an intensity image based on the image picked up in the image pickup step.

       (27.d.)   Acquiring range information from the picked-up image .

       (27.e.)   Performing geometric transformation for the intensity image based on the range

             information acquired in the range information acquisition step.

and

       (27.f.)   Extracting a difference between the geometric-transformed intensity image and an

             intensity image acquired in advance.

       (27.g.)   Storing, as the geometric transformed intensity image, an initial geometric-transformed

             image in a time-series transformed in the geometric transformation step.

       (27.h.)   Making comparison between successive geometric-transformed intensity images in the

             time-series transformed in the geometric transformation step

       (27.i.)   Retrieving only differential data between successive geometric-transformed intensity

                    images in the time-series as storage data and geometric-transformed intensity images

                    subsequently transformed in the time-series.

      (27.j.)   The stored geometric-transformed intensity image is the initial geometric-transformed

           .   intensity image and the differential data between successive geometric-transformed

                    images in the time-series.

73.   The rejection of Claim 17 follows similarly.

74.   *The following is in regard to Claims 18 and 28.* [Bunke99] suggests that the writing surface, or "image holding medium", is a manuscript sheet (i.e. a piece of paper).

75.   *The following is in regard to Claims 19 and 29.* As discussed above with regard to claim 27, the intensity image acquired in advance as a processing target in the image extracting step is a preceding frame image inputted precedent to the geometric transformation step. In fact, this limitation follows directly from the language of claim 27 ("...geometric-transformed intensity image and the intensity image acquired in advance").

76.   *The following is in regard to Claims 20 and 30.* Clearly, if one desires to process an image, or any data for that matter, acquired at some prior time, then the associated data must be stored in some means of storage, at the time of its acquisition (i.e. *in advance* of the processing). Therefore, the feature proposed in claim 30 would be inherent to the combination of [Bunke99] and [Saund98(383)].

77.   *The following is in regard to Claim 37.* Claim 37 recites substantially the same limitations as Claim 27. (The claimed storage medium contains a program that merely implements the method of Claim 27). Therefore, with regard to Claim 37, remarks analogous to those presented above relating to Claim 27 are applicable.

78.   *The following is in regard to Claims 38-40.* Note that, since capturing an image of a scene (the image holding medium) having projected light cast upon it is essentially the same as picking up the projected light, Claims 38-40 do not introduce anything substantively different from what is claimed in Claims 17, 27, and 37, respectively.

Therefore, with regard to Claims 38-40, remarks analogous to those presented above relating to Claims 17, 27, and

37 are applicable.

79.   Claims 21-24, 26, 31-34, and 36 are rejected under 35 U.S.C. § 103(a) as being unpatentable over

[Bunke99] and [Saund98(383)], as applied to Claims 17 and 27 above, in further view of [Stolfo98] (S. Stolfo, *U.S.*

Patent 5,748,780: Method And Apparatus For Imaging, Image Processing And Data Compression, Filing Date:

June 1994).

80.   *The following is in regard to Claims 21 and 31.* [Bunke99] and [Saund98(383)] do not expressly show or

suggest:

(31.a.)   Storing plural pieces of document format data in a document database.

(31.b.)   Performing matching between a geometric-transformed intensity image and the pieces of

document format data stored in the document database.

(31.c.)   Extracting a difference between the geometric-transformed intensity image and the

pieces of document format data stored in the document database[9].

81.   Many so-called "form drop-out" techniques perform these steps. For example, [Stolfo98] discloses a

methodology for extracting or separating handwritten text from standardized documents, such as various types of

checks ([Stolfo98], Summary of Invention, paragraph 1 and column 21, lines 5-11). According to this methodology:

(31.a$_{Stolfo}$.)   Plural pieces of document format data[10] are stored in a document database. This is

suggested in [Stolfo98] column 7, lines 23-25.

(31.b$_{Stolfo}$.)   A scanned (intensity) image is matched to the pieces of document format data stored in

the document database ([Stolfo98] column 7, lines 8-22).

(31.c$_{Stolfo}$.)   A subtraction (i.e. extraction of the difference) is performed between the intensity image

---

9   Note that step (31.c) essentially consists of (27.f) extracting difference between the geometric-transformed intensity image and the intensity
     image acquired in advance, where the latter intensity image would correspond to the document format data.

10   The (somewhat misleading) term document format data is taken in this case to mean a standard document image or template. This is
     consistent with the Applicant's usage of the term.

and the pieces of document format data stored in the document database ([Stolfo98]

column 7, lines 19-22).

82.   The teachings of [Bunke99], [Saund98(383)] and [Stolfo98] are combinable because they are analogous art.

Specifically, [Saund98(383)] demonstrate a method for scanning a document that eliminates skew. [Stolfo98], on the

other hand, discuss a method of handwriting extraction that calls for the document scanning. [Bunke99] similarly

disclose a handwriting extraction method. In particular, [Stolfo98]'s method calls for scanned documents to be,

among other things, skew-corrected (e.g. [Stolfo98] column 14, lines 62-66 and column 3, lines 29-43). Therefore, it

would have been obvious to one of ordinary skill in the art, at the time of the applicant's claimed invention, to apply

the platenless book scanning methodology of [Saund98(383)] to a handwriting extraction method such as that of

[Stolfo98], thereby generating a geometrically-transformed intensity image for which to process. The motivation for

doing so would have been to provide handwriting extraction from standardized documents that have been skew-

corrected. Again, assuming the geometric-transformation is precedent, one can conclude the combination of

[Bunke99], [Saund98(383)] and [Stolfo98] includes the following elements:

(31.a.)   Storing plural pieces of document format data in a document database.

(31.b.)   Performing matching between a geometric-transformed intensity image and the pieces of

document format data stored in the document database.

(31.c.)   Extracting a difference between the geometric-transformed intensity image and the

pieces of document format data stored in the document database.

83.   The rejection of Claim 21 follows similarly.

84.   *The following is in regard to Claims 22 and 32.* Both [Bunke99] (cf. [Bunke99] Section 3, paragraph 4) and

[Stolfo98] (e.g. [Stolfo89] column 4, lines 58-63) suggested applying handwriting recognition or optical character

recognition to the extracted handwriting.

85.   *The following is in regard to Claims 23-24 and 33-34.* According to [Stolfo89] ([Stolfo89] column 26, lines

66-67 to column 27, lines 1-6), the extracted handwriting, specifically the check signature, can be authenticated

against a database of signatures (i.e. a "handwriting history data of registered users in a authentication information

database"). Clearly, performing matching between the input image and the handwriting history data stored in the

authentication information database would be inherent to such an authentication process. Furthermore, given the

discussion above, it can be reasonably assumed that the geometric-transformed intensity image serves as input. In

addition, the authentication taught by [Stolfo89] is of the signature extracted from the scanned document.

86.   *The following is in regard to Claims 26 and 36.* It should be (overwhelmingly) apparent that the range

information would be stored after being derived. Furthermore, it should be apparent from [Saund98(383)] (e.g.

[Saund98(383)] Fig. 1) that the camera is in a fixed position relative to the document, or more particularly, the

document holder[11].

87.   Claims 25 and 35 are rejected under 35 U.S.C. § 103(a) as being unpatentable over [Bunke99] and

[Saund98(383)], as applied to Claims 17 and 27 above, in further view of [Wellner96] (P.D. Wellner, *U.S. Patent

5,511,148: Interactive Copying System*, Filing Date: April 1994).

88.   *The following is in regard to Claims 25 and 35.* Neither [Saund98(383)] nor [Bunke99] expressly show or

suggest displaying an image produced as a result of performing geometric transformation for the intensity image

based on the range information.

89.   [Wellner96] discloses an "interactive digital desktop", that is, a system for generating new documents from

originals containing text and/or images employing e.g. a camera-projector system focused on a work surface

([Wellner96] Abstract). In [Wellner96] (depicted in [Wellner96] Fig. 1) a video projector 8 is mounted adjacent the

camera 6 and projects onto the surface 2 a display 21 which is generally coincident with the field of view of the

camera 6, and which, in the example shown, includes an image of a newly created document 20. This projected

document image is skew-corrected ([Wellner96] column 10, lines 61-62). In this way, [Wellner96] teaches

displaying an image produced as a result of performing geometric transformation for the intensity image.

---

11   If it were not the various transforms (in particular the perspective transform) derived according to [Saund98(383)] would be invalid, or the entire system would have to be constantly recalibrated. One of ordinary skill in the art would easily recognize this.

90.   The teachings of [Wellner96] are combinable with those of [Saund98(383)] and [Bunke99] because they

are analogous art. Specifically, [Wellner96] concerns both platenless scanning (see [Wellner96] Fig. 1) and text

extraction and recognition ([Wellner96] column 9, lines 60-63). Therefore, it would have been obvious to one of

ordinary skill in the art, at the time of the applicant's claimed invention, to incorporate the projection or display

functionality of the [Wellner96] digital desktop into the combination of [Bunke99] and [Saund98(383)], wherein

skew-correction is performed according to the method proposed by [Saund98(383)]. The motivation for doing so

would have been to provide the user with visual feedback relating to the alignment of the document.

Citation of Relevant Prior Art

91.   The prior art made of record and not relied upon is considered pertinent to Applicant's disclosure:

92.   "Digital Desktop" applications that are responsive to a user's handwriting.

        [Arai95]   T. Arai et al., *U.S. Patent 5,436,639: Information Processing System*. Filing Date:

                        March 1994.

        [Kuzunuki98]   S. Kuzunuki et al., *U.S. Patent 5,732,227: Interactive Information Processing*

                        System Responsive to User Manipulation and Displayed Images. Filing Date: July

                        1995.

93.   Platenless document scanners. [Yamada04] removes hand regions on the scanned document and determines

difference images. [Matsui99] uses structured light.

        [Yamada04]   K. Yamada, *U.S. Patent 6,697,536: Document Image Scanning Apparatus And*

                        Method Thereof. Filing Date: April 2000.

        [Matsui99]   T. Matsui, *U.S. Patent 5,886,342: Image Reader And Document Curvature*

                        Measurement Using Visible And Infrared Light. Filing Date: April 1997.

94.   Devices for transcribing, into electronic form, markings drawn on whiteboad or blackboard. These devices
use video cameras and involve some form of geometric transformation.

[Saund96]   E. Saund, *U.S. Patent 5,528,290: Device For Transcribing Images On A Board Using*
             A Camera Based Board Scanner. Filing Date: September 1994.

[Suand02]   E. Saund, *U.S. Patent 6,411,732: Method For Interpreting Hand Drawn Diagrammatic*
             User Interface Commands. Filing Date: July 1997.

95.   [Iida01] discloses a method scanning bound documents. [Iida01] includes provisions for removing flesh
regions from the scanned documents.

[Iida01]   *U.S. Patent 6,256,411: Image Processing Device And Method For Detecting Objects In*
             Image Data. Filing Date: May 1998.

96.   Other literature from Perona et al.

[Munich96]   M.E. Munich and P. Perona, *Visual Input for Pen-Based Computers*, Proceeding of
             the ICPR'96 © IEEE, 1996.

[Munich98]   M.E. Munich and P. Perona, *Camera-Based ID Verification by Signature Tracking*,
             Proceedings of the 5th European Conference on Computer Vision-Volume, 1998.

97.   Works related to video-based, online handwriting tracking, extraction and recognition. [Yamasaki96] is
quite similar to [Bunke99]. [CrowleyVis95] and [CrowleyFin95] track a finger, as opposed to writing implement.
They employ difference images.

[Fink01]   G. Fink et al., *Video-Based On-Line Handwriting Recognition*, Proceedings of
                the International Conference on Document Analysis and Recognition © IEEE,
                2001.

[Wienecke01]   M. Wienecke et al., *A Handwriting Recognition System Based on Visual Input*,
                2nd International Workshop on Computer Vision Systems © IEEE, 2001.

[Yamasaki96]   T. Yamasaki et al., *A New Tablet System for Handwriting Characters and Drawing Based on the Image Processing* © IEEE, 1996

[CrowleyVis95]   J.L. Crowley et al., *Finger Tracking as an Input Device for Augmented Reality*, Proceedings of the IWAFGR, 1995.

[CrowleyFin95]   J.L. Crowley et al., *Vision for a Man Machine Interaction*, EHCI '95, August 1995.

Conclusion

98.   **THIS ACTION IS MADE FINAL.** Applicant is reminded of the extension of time policy as set forth in 37 CFR 1.136(a).

99.   A shortened statutory period for reply to this final action is set to expire THREE MONTHS from the mailing date of this action. In the event a first reply is filed within TWO MONTHS of the mailing date of this final action and the advisory action is not mailed until after the end of the THREE-MONTH shortened statutory period, then the shortened statutory period will expire on the date the advisory action is mailed, and any extension fee pursuant to 37 CFR 1.136(a) will be calculated from the mailing date of the advisory action. In no event, however, will the statutory period for reply expire later than SIX MONTHS from the mailing date of this final action.

Any inquiry concerning this communication or earlier communications from the examiner should be directed to Kevin Siangchin whose telephone number is (703)305-7569. The examiner can normally be reached on 9:00am - 5:30pm, Monday - Friday.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, Amelia Au can be reached on (703)308-6604. The fax phone number for the organization where this application or proceeding is assigned is 703-872-9306.

Information regarding the status of an application may be obtained from the Patent Application Information

Retrieval (PAIR) system. Status information for published applications may be obtained from either Private PAIR

or Public PAIR. Status information for unpublished applications is available through Private PAIR only. For more

information about the PAIR system, see http://pair-direct.uspto.gov. Should you have questions on access to the

Private PAIR system, contact the Electronic Business Center (EBC) at 866-217-9197 (toll-free).

Kevin Siangchin

Examiner
Art Unit 2623

ks – 2/17/05 12:53 PM

AMELIA M. AU
SUPERVISORY PATENT EXAMINER
TECHNOLOGY CENTER 2600